Respiratory Tract Dysbiosis is Associated With Worse Outcomes in Mechanically-Ventilated Patients

Online Supplement

Extended Methods sections:

Sample acquisition from critically-ill, mechanically-ventilated patients:

All samples were collected within 72hrs of initiation of invasive mechanical ventilation via endotracheal intubation. In 211 subjects, samples were collected on the same day that informed consent was obtained by their legally authorized representatives (LAR), whereas in 15 subjects samples were collected the day after informed consent was obtained (but still within 72hrs from initiation of mechanical ventilation). For 75 subjects, samples were obtained with a time waiver of consent (Institutional Review Board – IRB approval for collection of biospecimens under a minimal risk protocol with subsequent acquisition of informed consent from study participants or their LARs). In these 75 subjects, median number of days (range) from sample to informed consent acquisition was 3 (1-36) days. From each subject, we obtained baseline samples for microbiome analyses of the oral and lung communities and plasma samples for measurement of host-response biomarkers, with the following sample acquisition protocol (1):

Oral swabs: a sterile cotton swab was gently rubbed on the posterior oropharynx next to the endotracheal tube for ~ 5 secs and then capped in the collection tube, labeled and frozen to -80°C as soon as possible until sample processing.

Endotracheal aspirates (ETAs): Distal tracheal secretions were suctioned through a closed endotracheal tube suctioning system, similar to the ETAs obtained during routine clinical care for microbiologic culture studies. In cases where suctioning did not return adequate (>5ml) amount of ETAs, we instilled 5mL of sterile saline through the tubing system and then repeated suctioning. In cases where sterile saline was used for sampling, we also collected the left-over saline not used for sample collection (~ 5mL)

as a negative control to examine for procedural contamination of our samples. Samples were directly collected in sputum collection traps, labeled and frozen to -80°C as soon as possible until sample processing.

Plasma samples: At the same time of respiratory sample collections, we also collected blood samples (10mL) (through central venous or arterial access or phlebotomy) collected in sodium citrate anticoagulated tubes ("blue-top") provided that the patients' hemoglobin concentration was >8g/dl to avoid risks of iatrogenic anemia. Samples were then transferred to the research laboratory for same day centrifugation and plasma collection and storage to -80°C.

Clinical data recording:

For each patient, we recorded detailed data abstracted from the electronic medical record (EMR) that were used in downstream analyses. Demographic and medical history data were collected from the EMR on the day of enrollment. Vital signs, laboratory results, mechanical ventilation parameters and chest radiography within 24hrs from enrollment were reviewed and the physiologically worse values (e.g. lowest blood pressure, or highest creatinine value) were recorded.

Demographic information: age, sex, height, weight, body mass index.

Pertinent medical history data: history of diabetes, chronic obstructive lung disease, congestive heart failure, chronic kidney disease, active neoplasm, chronic liver disease, solid organ transplant, immunosuppression (defined as use of chronic steroids, alkylating agents, antimetabolites, calcineurin inhibitors, mycophenolate, biologics, active chemotherapy against solid tumor or hematologic malignancy, or diagnosis of primary

immunodeficiency, such as common variable immunodeficiency, chronic granulomatous disease etc).

Vital signs: temperature, heart rate, systolic blood pressure (lowest), mean arterial pressure (lowest).

Laboratory results: white blood cell count, hemoglobin, platelets, carbon dioxide serum concentration, glucose, arterial pHa, partial pressure of oxygen concentration, partial pressure of carbon dioxide concentration.

Mechanical ventilation parameters: respiratory rate, tidal volume per kilogram ideal body weight, positive end-expiratory pressure, plateau pressure.

Chest radiography: presence of bilateral infiltrates, abnormal chest x-ray at enrollment.

Clinical microbiology results: We considered clinical microbiologic results from samples obtained within 48hrs of research sample timing acquisition so that such samples would be reflective of the same infectious process being studied by next-generation sequencing. Clinical cultures were obtained at the discretion of the treating physicians who were not involved in the research study. We considered microbiologic cultures of respiratory specimens (sputum, ETA or bronchoalveolar lavage - BAL) as positive when pathogenic bacterial species had been isolated by the clinical laboratory. Patients without positive cultures were considered those in whom respiratory specimen cultures were obtained and resulted as negative (i.e. with no growth or only normal respiratory flora detected, n=152) as well as those patients for whom no cultures were obtained at the decision of the treating physicians (n=76).

Severity of illness scores: sequential organ failure assessment (SOFA) score (2) (calculation does not include the neurologic component of SOFA score because all

patients were intubated and were receiving sedative medications, which impaired our ability to perform assessment of the Glasgow Coma Scale in a consistent and reproducible fashion). SOFA scores were calculated using the physiologically worse values within 24hrs of enrollment. Lung Injury Prediction Scores (LIPS) were calculated from baseline variables (3).

Medication administration: We recorded antibiotics and vasopressors administered during the first week of ICU course from intubation, as well as whether systemic antibiotics had been administered prior to ICU admission within the preceding 30 days (categorical variable). For systemic antibiotics administered in the ICU prior to the microbiota sampling, we performed detailed quantitative modeling of this antibiotic exposure by utilizing a published model that controls for i) dosing duration, ii) timing of administration relative to sample collection, and iii) antibiotic type and route of administration. For this **ICU antibiotic exposure score**, we utilized weighting scores for each specific antibiotic for building the antibiotic exposure model, and we also classified each antibiotic into broader categories of anti-bacterial spectra for descriptive purposes (Table E1). Given that we focused exclusively on profiling bacterial community profiles by 16S sequencing, we did not assess for the potential impact of antifungal or antiviral medications on non-bacterial microbiota.

Table E1. Weighting score and anti-bacterial spectrum classification for antibiotics administered in the ICU prior to sampling of respiratory microbiota.

Antibiotic	Weighting score	Anti-bacterial spectrum category
Amikacin	0.33	Gram-negative
Amoxicillin-clavulanate	0.5	Broad-spectrum
Ampicillin-sulbactam	0.5	Broad-spectrum
Azithromycin	0.17	Atypicals
Cefazolin	0.33	Gram-positive
Cefepime	0.5	Broad-spectrum
Cefotaxime	0.33	Broad-spectrum
Ceftazidime	0.33	Broad-spectrum
Ceftriaxone	0.33	Broad-spectrum
Cefuroxime	0.5	Broad-spectrum
Cephalexin	0.33	Gram-positive
Ciprofloxacin	0.5	Broad-spectrum
Clarithromycin	0.17	Atypicals
Clindamycin	0.17	Anaerobes
Erythromycin	0.17	Atypicals
Gentamicin	0.33	Gram-negative
Imipenem-cilastatin	0.5	Broad-spectrum
Levofloxacin	0.5	Broad-spectrum
Linezolid	0.33	Gram-positive
Meropenem	0.5	Broad-spectrum
Metronidazole	0.17	Anaerobes
Nafcillin	0.5	Gram-positive
Piperacillin-tazobactam	0.33	Broad-spectrum

Tobramycin	0.17	Gram-negative
Trimethoprim-sulfamethoxazole	0.33	Gram-positive
Vancomycin	0.33	Gram-positive

Outcomes: length of ICU stay, ventilator-free days (VFD), cumulative 30-day mortality, 30-day survival and time to liberation from mechanical ventilation.

Samples from healthy volunteers.

For comparisons of ecological metrics (alpha and beta-diversity) of the sequenced communities of critically-ill patients against communities reflective of the normal upper and lower respiratory tract microbiome, we used existing data from a previous study of respiratory microbiota in healthy volunteers. We processed available sequencing data (see Analytics section) from 23 healthy volunteers (HIV-negative adult subjects) that had participated in the Lung HIV Microbiome study (4). These subjects contributed 23 oral wash samples, 21 induced sputum samples and 23 bronchoalveolar lavage (BAL) samples. We compared the oral wash samples with oral swabs from our critically-ill patients (upper respiratory tract comparison), and then the induced sputum and BAL samples against the endotracheal aspirates of our critically-ill patients (lower respiratory tract comparison).

Laboratory methods:

Positive and Negative experimental control samples:

We considered three types of experimental negative control samples used to assess for possible contamination events from sample collection to DNA extraction and PCR amplification:

Sampling controls:

A. Critically-ill patients:

1. **Endotracheal aspirate controls**: In cases where sterile saline was used for sampling of ETAs, we collected the left-over saline not used for sample collection (~ 5mL) as a negative control to examine for contamination of the ETA from bacterial populations present in what is clinically considered sterile saline syringes. We extracted DNA from these saline samples and performed PCR amplification and 16S rRNA gene sequencing as done for the clinical samples.

B. Healthy volunteers;

1. **BAL negative controls**: Before bronchoscopy, 10 to 50 mL of sterile 0.9% saline were washed through the bronchoscope and collected as a control for DNA in the bronchoscope
2. **Oral wash negative controls**: a sterile saline in a sample collection cup from the saline used to perform the oral wash was used as negative control.

Experimental processing controls:

1. **DNA extraction negative controls**: We added DNA-free sterile water in one extraction column per each batch of clinical sample DNA extraction (~12-15 samples per experiment).
2. **PCR-negative controls**: We added DNA-free sterile water in the PCR reaction mix in amount equal to the amount of template DNA from clinical samples used (typically 6 microliters).

3. **PCR amplification positive controls**: We utilized the ZymoBIOMICS Microbial Community DNA Standard (Zymo Research, Irvine, CA), a mock microbial community consisting of genomic DNA of eight bacterial strains. We utilized 1 microliter of the genomic mixture with concentration of 10ng/microliter for each reaction.

DNA extractions: We extracted bacterial DNA directly from samples (oral swabs and endotracheal aspirates - ETAs) using the Powersoil (MoBio) extraction kit following the manufacturer's instructions, as previously described (4). Due to high viscosity of ETA samples, we pretreated them with Dithiotreitol (0.1% DTT in phosphate buffered saline) in 1:1 dilution in order to dissolve the mucus and allow usability in DNA extraction columns.

16S rRNA gene sequencing: We amplified extracted DNA by PCR using the method of Caporaso et al.(5) and the Q5 HS High-Fidelity polymerase (NEB) targeting the V4 hypervariable region of the 16S rRNA gene. We utilized reagent controls for each step of the process (DNA extraction and PCR amplification). We amplified four microliters per reaction of each sample with a single barcode in triplicate 25 microliter reactions. Cycle conditions were 98˚C for 30s, then 33 cycles of 98˚C for 10s, 57˚C for 30s, 72˚C for 30s, with a final extension step of 72˚C for 2 min. We combined triplicates and purified with the AMPure XP beads (Beckman) at a 0.8:1 ratio (beads:DNA) to remove primer-dimers. We quantitated eluted DNA on a Qubit fluorimeter (Life Technologies). We performed sample pooling on ice by combining 20 ng of each purified band. For negative controls and poorly performing samples, we used 20 microliters of each sample. We purified the sample pool with the MinElute PCR purification kit. The final sample pool underwent two more

purifications – AMPure XP beads to 0.8:1 to remove all traces of primer dimers and a final cleanup in Purelink PCR Purification Kit (Life Technologies). We quantitated the purified pool in triplicate on the Qubit fluorimeter prior to preparing for sequencing. We prepared the sequencing pool according to instructions by Illumina, with an added incubation at 95˚C for 2 minutes immediately following the initial dilution to 20 picomolar. We then diluted the sequencing pool to a final concentration of 7 pM + 15% PhiX control. Amplicons were sequenced on the Miseq platform.

16S rRNA gene quantitative PCR (qPCR): We performed qPCR reactions of the V3-V4 region of the 16S rRNA gene to quantify the bacterial load in each sample, with the BactQuant protocol (6). PCR reactions were prepared in a total volume of 20µL, which consisted of 2µL of 10XPCR buffer, 3.5 mmol/L $MgCl_2$, 0.2 mmol/L deoxynucleoside triphosphate, 0.5 µmol/L forward and reverse primers, 0.225 µmol/L probe, 0.75 U of Platinum Taq polymerase (Invitrogen), and 2µL of each DNA. The forward and reverse primers and the probe sequences that were used to amplify DNA templates encoding the V3-V4 region of the 16S rRNA gene were identical to those previously described (6). The DNA was amplified in duplicate, and mean values were calculated. A standard curve was created from serial dilutions of plasmid DNA containing known copy numbers of the template. The assays were performed on the LightCyler System (Roche) using the following PCR conditions: 95°C for 5 min, followed by 50 cycles at 95°C for 15 s and at 60°C for 1 min.

Analytics:

16S Sequence quality control:

Sequences from the pooled sequencing run were demultiplexed into individual sample/replicate fastq files. Each fastq file was then processed through the Center for Medicine and the Microbiome (CMM) custom modular read QC pipeline that was configured to perform the following steps: low complexity filtering, QV trimming, Illumina sequencing adapter trimming, and 16S primer trimming. Low complexity filtering utilized NCBI BLAST's dustmasker (7). Reads with greater than 80% low complexity regions were filtered. QV trimming and filtering utilized the FASTX Toolkit. The trimming threshold was QV>25 from the 3' end, and length >125 bp after trimming. The subsequent filtering threshold was >25 across >95% of the read. The Illumina sequencing adapter and 16S primer trimming was performed with cutadapt (8).

16S Clustering and annotation**:**

Paired sequences with forward and reverse reads passing the Quality Control filtering and trimming steps were then mated (end aligned and a consensus sequence computed) using the make.contigs function of Mothur. Consensus sequences were screened to limit the overlap mismatch to no more than 20%. The maximum number of N's allowed in the overlap was 4 and the minimum overlap was required to be greater than 25bp. Consensus sequences passing screening were then passed through CMM's 16S clustering and annotation pipeline, a Mothur-dependent wrapper designed to streamline and automate the execution of the following Mothur steps in version v.1.39.1: unique.seqs, align.seqs, screen.seqs, filter.seqs, second uniq.seqs, pre.cluster, chimera.uchime, remove.seqs, classify.seqs, dist.seqs, cluster, make.shared, and classify.otu. Mothur output files were then reformatted to sample x category (taxonomic levels or operational taxonomic units [OTU] at 97% sequence similarity). For downstream

statistical modeling and analyses, we utilized taxonomic tables at the genus level, with taxonomic assignments performed using a naïve Bayes k-mer classifier in conjunction with the Ribosomal Database Project (RDP) 16S rRNA gene sequences (8).

Taxa table edits:

The taxa table was filtered for low abundance taxa (relative abundance <0.005%) and singletons. From a total of 6,239,912 reads, filtering for such low abundance taxa resulted in a total of 6,214,752 reads (99.6% of initial read count) and 278 unique genera for analyses. We did not filter clinical samples for any taxa detected in the negative control samples. Clinical samples with very few reads (<200) were excluded from analyses (7 ETA samples and 2 oral swabs).

Statistical Analyses:

We calculated descriptive statistics of baseline characteristics and performed nonparametric comparisons using the R software (v.3.5.1). Biomarker values were log-transformed for analyses. With logistic regression models combining biomarker and clinical variables, patients were assigned to hyper- vs. hypo-inflammatory subphenotypes as previously described (9). These logistic regression models were derived by application of latent class analyses for derivation of subphenotypes in patients with ARDS and at-risk for ARDS, followed by probabilistic graphical models for input variable selection (feature selection). In patients with ARDS, the following logistic regression model was used (with demonstrated accuracy against latent class analysis 93.2%):

$$Subphenotype = 45.2 + 23.0*(pHa) + 13.2*(Temperature) - 144.4*(TNFR1) - 78.0*(Creatinine) + 33.9*(CO_2) - 24.0*(RAGE) - 59.9*(Fractalkine).$$

In patients without ARDS, the following logistic regression model was used (with demonstrated accuracy against latent class analysis 98%):

Subphenotype = 104.5 + 23.0(pHa) + 34.9*(Temperature) – 171.9*(TNFR1) – 84.3*(IL-10) – 64.5*(Angiopoetin-2) – 67.1*(Fractalkine).*

For ecological analyses of alpha diversity (Shannon index) and beta diversity (Manhattan distances with permutational analysis of variance [permanova] at 1000 permutations), we used the R *vegan* package and visualized beta-diversity differences with principal coordinates analyses (PCoA) plots.

To agnostically examine for distinct clusters of microbial composition ("meta-communities") in our samples, we applied unsupervised Dirichlet Multinomial Models (DMM) with Laplace approximations to define the optimal number of clusters in our dataset (10). The DMM approach has been shown to outperform traditional Gaussian multivariate techniques for deriving distinct meta-communities, i.e. dividing sample pools into subgroups of communities with similar compositions. DMM clusters were derived in the trimmed dataset following exclusion of low abundance taxa (i.e. those with <0.005% relative abundance). To examine the robustness of DMM clustering, we performed 10-fold cross validation, by dividing the sample pool into 10 random subgroups of similar size, then deriving the DMM clusters based on 9 subgroups and applying the prediction on the left-out subgroup iteratively until a new prediction was obtained for each sample. Then, we checked concordance of the original clustering result from the original complete cohort against the predictions from the 10-fold cross validation. This iterative process demonstrated that 86% of samples were predicted to belong to their originally assigned cluster, thus demonstrating robustness of DMM clustering.

We undertook two parallel and complementary approaches for examining associations between microbiome profiles and clinical outcomes.

A. Unsupervised approach: We examined for associations between the agnostically derived DMM clusters with host-response subphenotypes and clinical outcomes (30 day survival and time-to-liberation from mechanical ventilation) in regression and cox-proportional hazards models, respectively. For time-to-event analyses (30 day survival and time-to-liberation), time zero was defined as the time of intubation. Cox-proportional hazards models for the independent effects of DMM clusters were adjusted for potential clinical confounders that were differentially distributed between clusters. For ETA clusters, cox models were adjusted for age, history of COPD, diagnosis of ARDS, extra-pulmonary sepsis, antibiotics prior to ICU admission (categorical variable) and ICU antibiotic exposure score (numerical score). For oral swab clusters, cox models were adjusted for age, history of COPD, history of immunosuppression, antibiotics prior to ICU admission (categorical variable) and ICU antibiotic exposure score (numerical score). Logistic regression models for host-response subphenotypes were adjusted for the same variables, with the exception of the model for ETA clusters, in which we excluded the variables of ARDS and extra-pulmonary sepsis because variables are considered to be part of the causal pathway between lung microbiota, clinical ARDS/sepsis and host-response biomarkers. Following cluster-level analyses, we then pursued genus-level analyses to delineate the contributions of individual genera in outcome prediction. We performed

relative abundance transformations with the additive log ratio for the top 10 genera in each DMM cluster and examined for associations with outcomes (hyperinflammatory subphenotype, 30-day Mortality and VFDs), adjusted for covariates and multiple-testing with the Benjamini-Hochberg method.

B. Supervised approach: We stratified our cohort by observed clinical outcomes, i.e. in survivors vs. non-survivors and in tertiles of VFDs (0-13, 14-23, >23 days). We then examined for differences in alpha and beta diversity, bacterial load as well as individual genera relative abundance (additive log-ratio transformed), independently for oral swabs and ETAs. Individual genera associations were adjusted for multiple-testing with the Benjamini-Hochberg method.

In a subsequent integrative step, we utilized the predictive features highlighted by both unsupervised and supervised methods (namely alpha diversity and specific genera abundance) for the development of a simple and generalizable predictive index. To that end, we aimed to define upper and lower respiratory tract communities with profiles similar to the normal oral and lung microbiome and no evidence of dysbiosis. Based on our empirical observations with the DMM clusters and the individual genera associations, we identified that the high alpha diversity cluster 1 in oral swabs and cluster 3 had the best outcomes, and that high relative abundance of oral-origin bacteria (e.g. *Prevotella_7, Streptococcus, Veillonella*) was associated with lower mortality and/or more VFDs. We identified partially overlapping sets of oral-origin bacteria in ETA and oral swab samples (protective bacteria). We then utilized independent receiver-operator characteristic (ROC) curves to define optimal thresholds of protective bacteria relative abundance in ETA and

oral swab samples for prediction of mortality. Additionally, we defined a Shannon threshold that separated the prognostically favorable cluster 3 in ETA and cluster 1 in oral swabs as 1.98 based on the distributions of Shannon index by cluster. Following determination of these thresholds, we then combined the Shannon threshold (≥1.98) and the protective bacteria relative abundance (≥30% for ETAs and ≥70% for oral swabs) into a simple **Dysbiosis Index**: samples that met both thresholds were considered to have no evidence of dysbiosis, whereas samples that did not meet both thresholds were considered to have evidence of dysbiosis. Then, we stratified our cohort by the Dysbiosis Index and examined for associations with outcomes (hyperinflammatory subphenotype, 30-day survival and time-to-liberation).

Figure E1. Number of reads (high quality 16S rRNA gene sequences) by sample type. Clinical samples are separated from experimental control samples by a dashed line. Clinical samples from ICU patients or healthy volunteers had much higher number of 16S reads compared to experimental control samples ($p<0.001$). Seven ETA samples and 2 oral swabs had low number of reads (<200) and were excluded from further analyses.

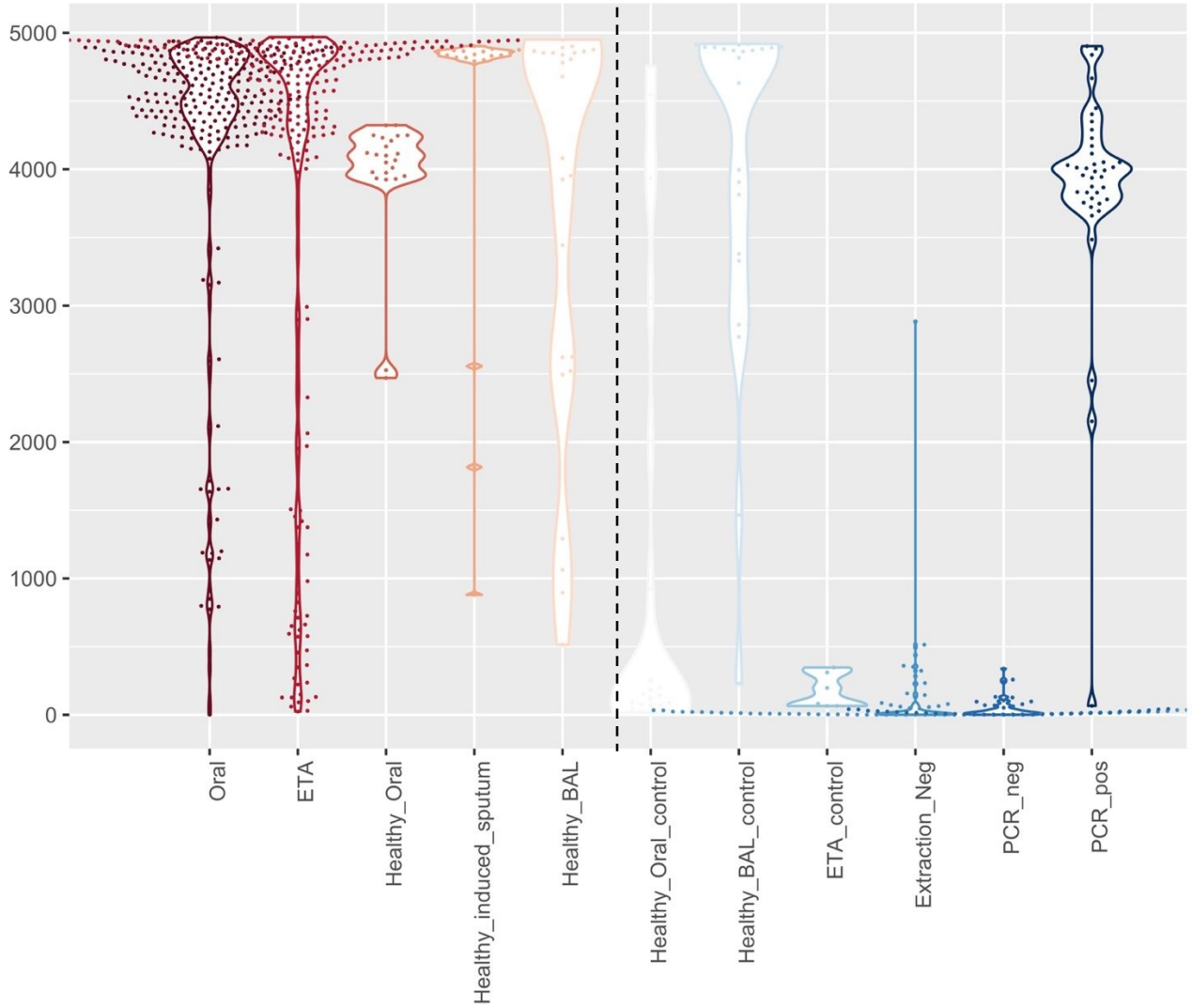


Figure E2. Upper and lower respiratory tract samples from critically-ill patients have lower alpha diversity and significantly different taxonomic composition from healthy controls. A. Alpha diversity (Shannon Index). B. Principal Coordinates Analysis plot for visualization of beta-diversity (Manhattan distances) between oral swabs and endotracheal aspirates (ETA) from intensive care unit (ICU) patients and healthy control upper (oral wash) and lower respiratory tract (bronchoalveolar lavage – BAL or induced sputum) samples. Diversity metrics were derived at the genus level. **** indicates p-value <0.0001.

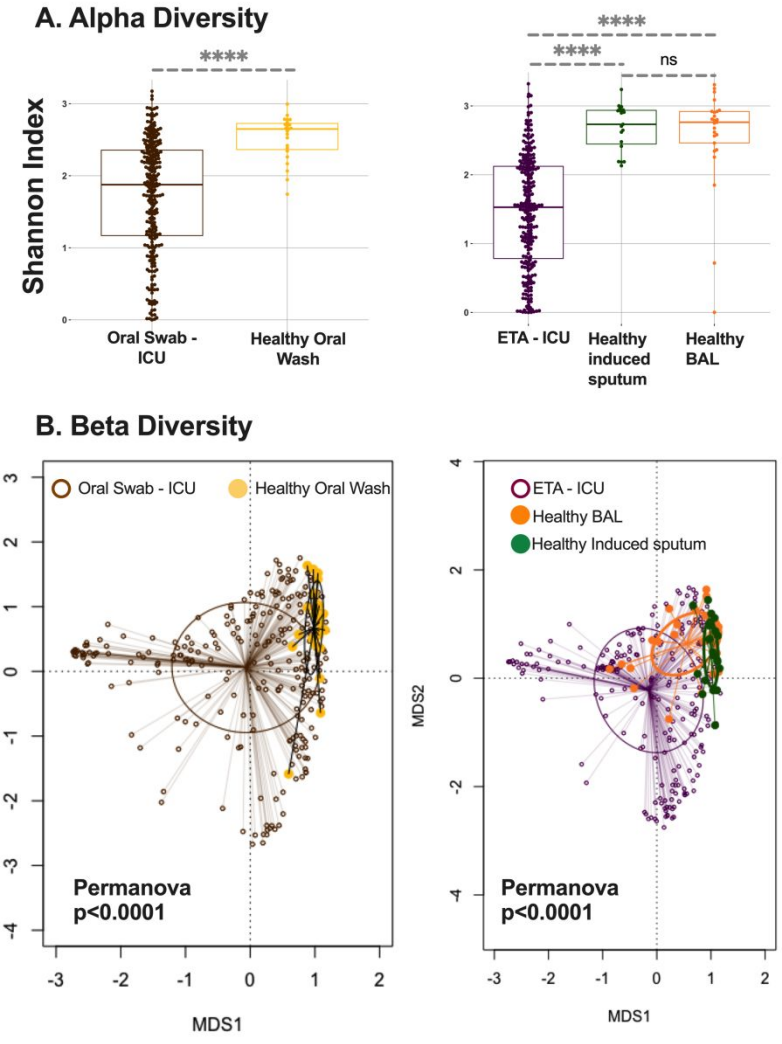


Figure E3. Oral swabs had ~20-fold higher bacterial load compared to endotracheal aspirates. Log-transformed number of copies of the 16S rRNA gene per microliter of extracted DNA solution by quantitative polymerase chain reaction (qPCR) are shown on the y-axis. **** indicates p-value <0.0001. Of note, 17 ETA samples (11%) had poor performance by qPCR (undetectable number of 16S gene copies). By Illumina sequencing, only one of these 17 samples had low performance (<200 number of reads). In the remaining samples, we detected significant correlation between number of qPCR gene copies and Illumina number of reads ($r = 0.37$, $p<10^{-5}$). Procedural variability with utilization of a ~5ml saline flush during sample acquisition may account for some of the observed variation in ETA qPCR load, resulting in dilution of certain samples. However, the observed variation in qPCR load (~3000 fold difference in load across the interquartile range of ETAs shown by the borders of the green boxplot) suggests that inter-sample variation from actual bacterial load present outweighs potential variability introduced from sample dilution.

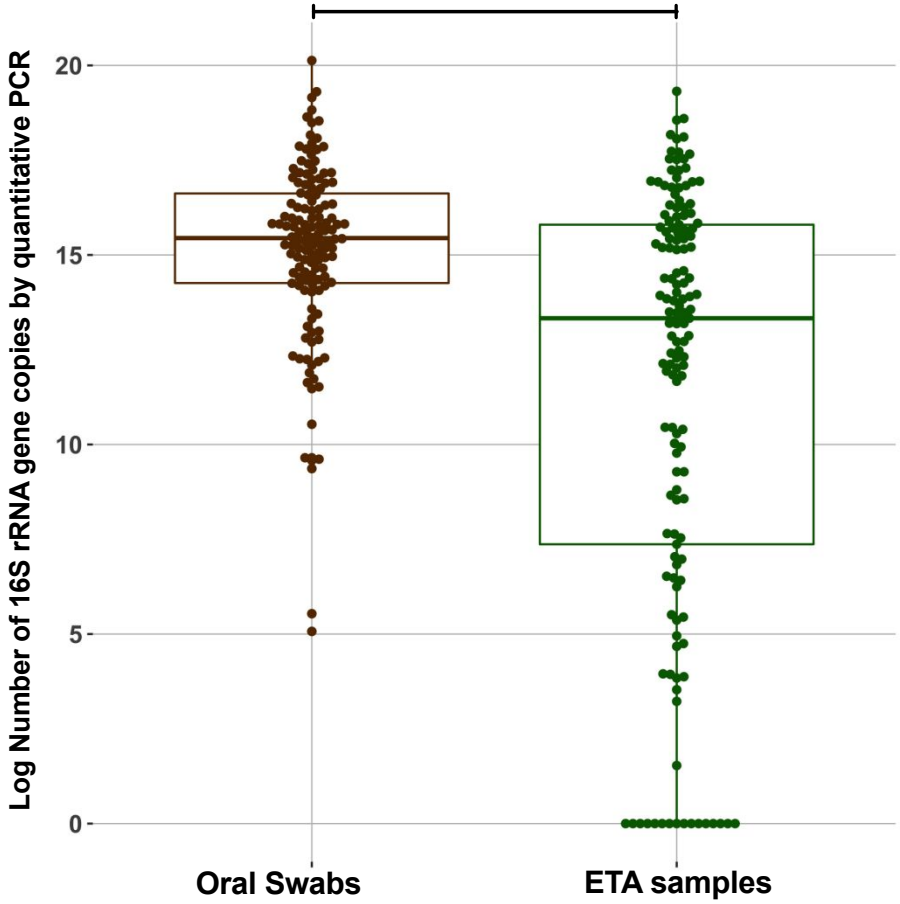

Log Number of 16S rRNA gene copies by quantitative PCR
20
15
10
5
0
Oral Swabs
ETA samples

Figure E4. Time interval from sample acquisition from intubation did not have a measurable impact on microbial community profile differences in oral swabs (left panel) or endotracheal aspirate samples (right panel). We stratified our 72hr sampling window from intubation in three intervals (0-24hrs, 24-48hrs, 48-72hrs). We did not identify any significant differences in alpha diversity (panel A), bacterial load by quantitative PCR (panel B) or beta diversity differences (panel C).

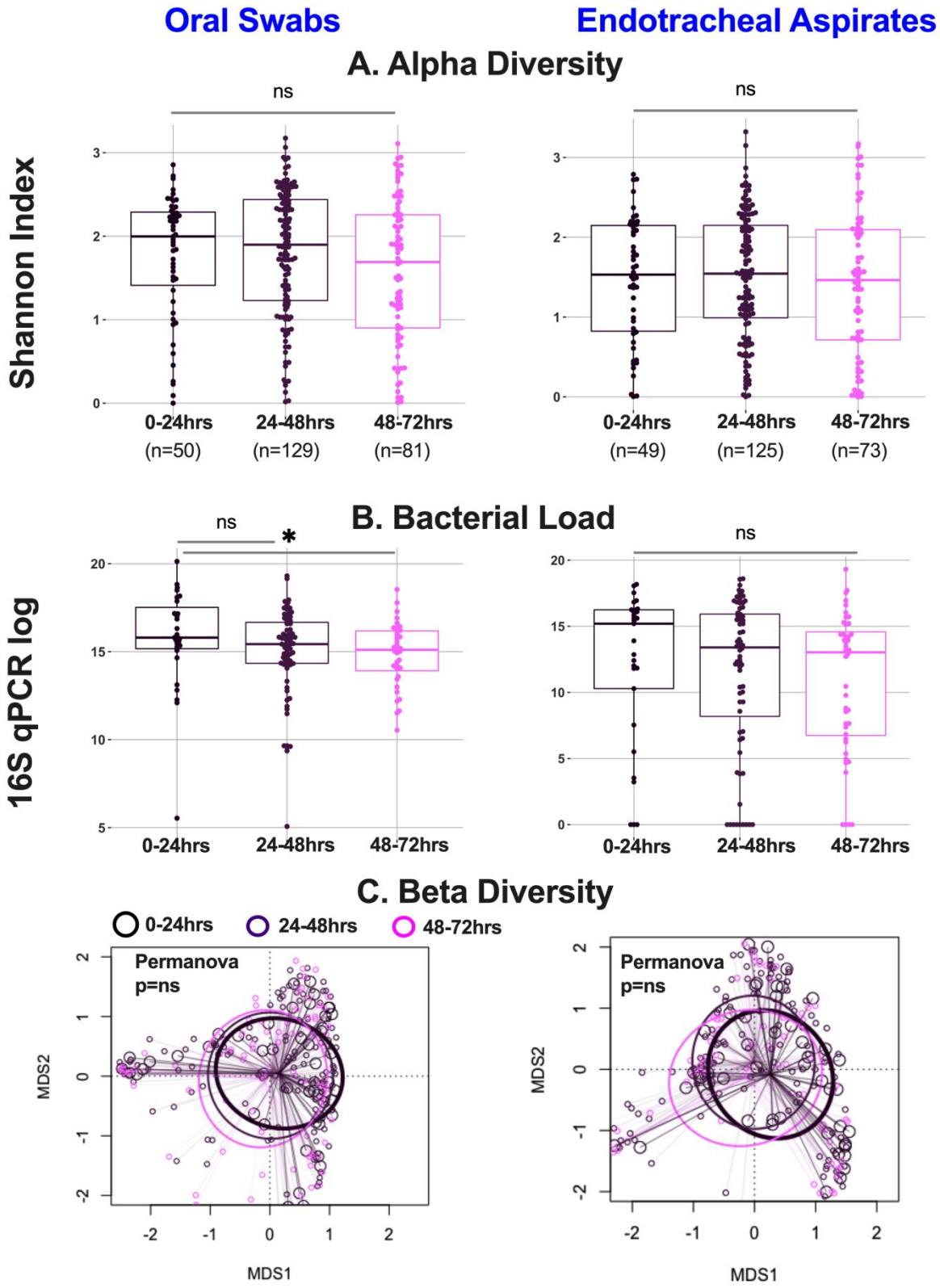

Oral Swabs
Endotracheal Aspirates
A. Alpha Diversity
Shannon Index
ns
ns
0-24hrs
(n=50)
24-48hrs
(n=129)
48-72hrs
(n=81)
0-24hrs
(n=49)
24-48hrs
(n=125)
48-72hrs
(n=73)
B. Bacterial Load
16S qPCR log
ns
*
ns
0-24hrs
24-48hrs
48-72hrs
C. Beta Diversity
0-24hrs
24-48hrs
48-72hrs
Permanova
p=ns
Permanova
p=ns
MDS1
MDS2

Figure E5: Laplace approximation model fitting showed that optimal fit was offered by a 2 cluster model in the oral swabs and a 3 cluster model in the endotracheal aspirate (ETA) samples.

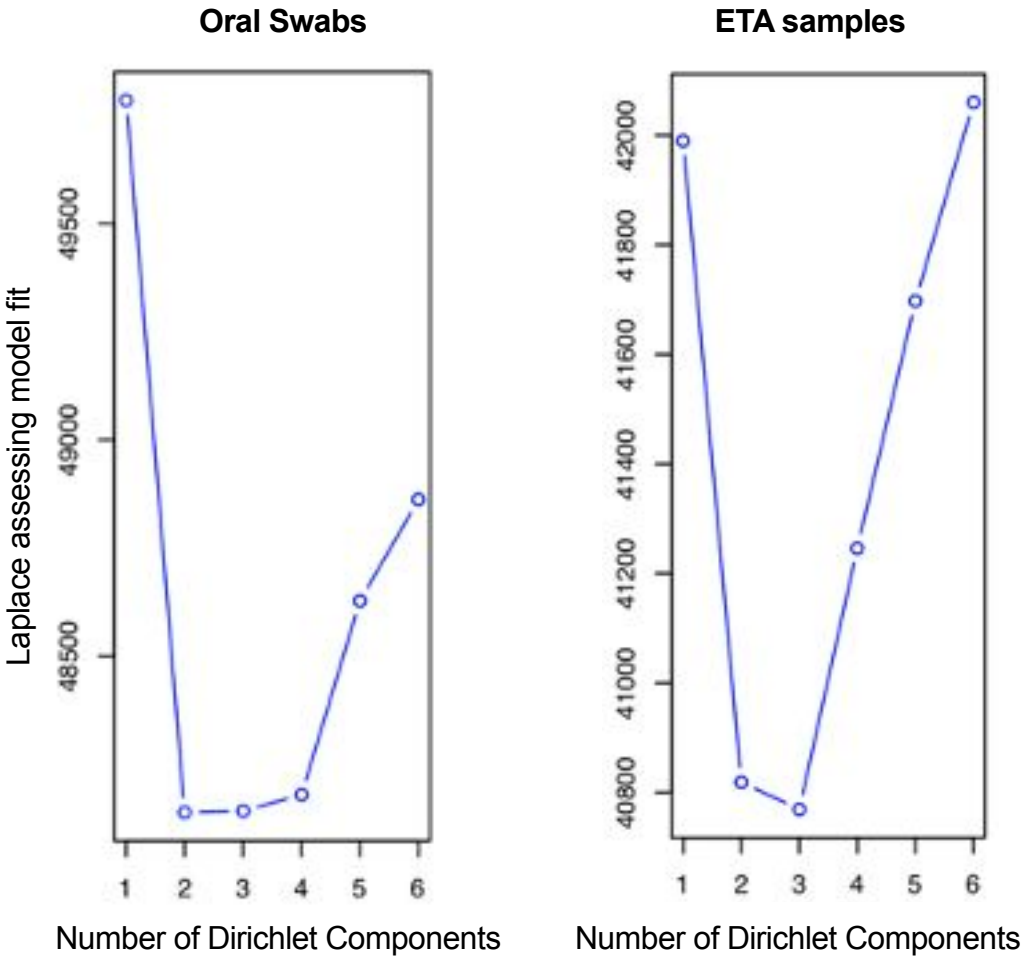


Table E2: Baseline characteristics and clinical outcomes by Dirichlet Multinomial Model clusters for oral swabs. Data are presented as median (with interquartile ranges) for continuous variables and N (%) for categorical variables. P-values for comparisons between clusters obtained from Wilcoxon test for continuous variables and Fisher's test for categorical variables. Statistically significant p-values (p<0.05) are highlighted in bold.

Variable	Cluster 1 (n=136)	Cluster 2 (n=103)	p-value
Age, median [IQR], years	56.9 [42.3, 65.4]	60.3 [51.3, 69.3]	**0.02**
Males, N (%)	73 (53.7)	51 (49.5)	0.61
BMI, median [IQR]	30.1 [25.5, 36.8]	28.4 [24.8, 34.9]	0.27
Diabetes, N (%)	46 (33.8)	38 (36.9)	0.72
COPD, N (%)	24 (17.6)	39 (37.9)	**<0.01**
Immunosuppression, N (%)	20 (14.7)	33 (32.0)	**<0.01**
ARDS, N (N%)	27 (19.9)	31 (30.1)	0.07
Pneumonia, N (%)	50 (36.8)	42 (40.8)	0.62
Extrapulmonary Sepsis, N (%)	26 (19.1)	19 (18.4)	1
Aspiration, N (%)	27 (19.9)	16 (15.5)	0.49
LIPS score, mean (SD)	5.5 [4.0, 7.0]	5.0 [4.0, 6.5]	0.46
SOFA score, median [IQR][a]	5.5 [4.0, 9.0]	6.0 [4.0, 8.0]	0.48
PaO2:FIO2 ratio, median [IQR], mmHg	177.5 [133.0, 205.0]	158.0 [109.5, 221.5]	0.18
WBC, median [IQR], x 10^9 per liter	12.9 [9.6, 16.4]	11.0 [7.9, 16.4]	0.07
Plateau pressure, median [IQR], cm	19.0 [16.5, 24.0]	22.0 [17.0, 26.0]	0.12
Positive respiratory cultures, N (%)	32 (23.5)	28 (27.2)	0.62
Systemic Antibiotics prior to ICU admission, N (%)	29 (21.3)	38 (36.9)	**0.01**
Antibiotics during ICU admission prior to sampling, N (%)	111 (81.6)	93 (90.3)	0.09

Systemic Steroids, N (%)	38 (28.1)	50 (48.5)	<0.01
Hyperinflammatory subphenotype, N (%)	31 (22.8)	20 (19.4)	0.64

Table E3. Anti-bacterial spectra of antibiotics used in ICU patients prior to microbiota sampling by Dirichlet Multinomial Models clusters in endotracheal sample specimens and in oral swabs. P-values are derived from non-parametric tests.

	Endotracheal Aspirate clusters			
Anti-bacterial spectrum	Cluster 1 (n=90)	Cluster 2 (n=78)	Cluster 3 (n=62)	p-value
Gram-positive, N (%)	59 (65.6)	67 (83.8)	44 (69.8)	**0.02**
Gram-negative, N (%)	3 (3.3)	2 (2.5)	0 (0.0)	0.36
Anaerobes, N (%)	7 (7.8)	4 (5.0)	6 (9.5)	0.57
Atypicals, N (%)	34 (37.8)	26 (32.5)	17 (27.0)	0.37
Broad-spectrum, N (%)	63 (70.0)	69 (86.2)	49 (77.8)	**0.04**
	Oral Swab clusters			
Anti-bacterial spectrum	Cluster 1 (n=136)	Cluster 2 (n=103)	p-value	
Gram-positive, N (%)	95 (69.9)	83 (80.6)	0.08	
Gram-negative, N (%)	2 (1.5)	4 (3.9)	0.45	
Anaerobes, N (%)	13 (9.6)	7 (6.8)	0.60	
Atypicals, N (%)	40 (29.4)	34 (33.0)	0.65	
Broad-spectrum, N (%)	101 (74.3)	85 (82.5)	0.17	

Table E4. Unadjusted clinical outcomes Dirichlet Multinomial Models clusters in endotracheal sample specimens and in oral swabs. P-values are derived from non-parametric tests.

	Endotracheal Aspirate clusters			
Outcome	Cluster 1 (n=90)	Cluster 2 (n=78)	Cluster 3 (n=62)	p-value
ICU length of stay, median [IQR], days	7.0 [4.0, 11.0]	11.0 [7.0, 17.0]	9.0 [5.0, 11.0]	**<0.01**
VFD, median [IQR], days	22.0 [7.2, 25.0]	15.5 [0.0, 23.0]	22.5 [17.2, 24.0]	**<0.01**
30 Day mortality, N (%)	21 (23.3)	25 (32.1)	6 (9.7)	**0.01**
	Oral Swab clusters			
Outcome	Cluster 1 (n=136)	Cluster 2 (n=103)	p-value	
ICU length of stay, median [IQR], days	8.0 [5.0, 11.0]	9.0 [5.0, 15.5]	**0.03**	
VFD, median [IQR], days	22.0 [12.0, 25.0]	19.0 [0.0, 23.0]	**0.01**	
30 Day mortality, N (%)	24 (17.6)	31 (30.1)	**0.03**	

Table E5: Logistic regression models for the association between cluster 2 membership and hyperinflammatory subphenotype classification in endotracheal aspirate (ETA) samples, oral swabs and in combined ETA and oral swab samples. Adjustments were performed for clinical variables differentially distributed between clusters. For the association between ETA clusters and the hyperinflammatory subphenotype, we excluded ARDS and extrapulmonary sepsis from the multivariate logistic regression models because both variables are considered to be in the causal pathway linking microbiota and host inflammation. Inclusion of bacterial load by qPCR in these models was not statistically significant and did not impact the association between DMM clusters and hyperinflammatory subphenotype.

	Hyperinflammatory subphenotype	**p-value**
ETA samples		
Cluster 2	**1.2 [1.1-1.9]**	**0.03**
Age	1.0 [1.0-1.0]	0.63
History of COPD	0.9 [0.8-1.0]	0.07
Antibiotics prior to ICU admission	1.0 [0.9-1.1]	0.94
Antibiotic Exposure Score	0.7 [0.2-2.4]	0.54
Oral swab samples		
Cluster 2	1.0 [0.9-1.1]	0.99
Age	1.0 [1.0-1.0]	0.53
History of COPD	0.9 [0.8-1.1]	0.14
Immunosuppression	1.1 [0.9-1.3]	0.78
Antibiotics prior to ICU admission	1.0 [0.9-1.2]	0.76
Antibiotic Exposure Score	0.8 [0.2-3.2]	0.81
Combined ETA samples and oral swabs		
Cluster 2 in both ETA and oral swabs	1.1 [0.9-1.3]	0.18
Age	1.0 [1.0-1.0]	0.44
History of COPD	0.9 [0.8-1.1]	0.39
Antibiotics prior to ICU admission	1.0 [0.8-1.1]	0.59

Antibiotic Exposure Score	0.6 [0.2-2.2]	0.42

Figure E6. Taxonomic bar plots for individual endotracheal aspirate samples, stratified by Dirichlet Multinomial Models clusters. Taxonomic composition is shown as stacked bar-graphs, with each bar representing a patient's community, with taxa colored individually and heights of component bars corresponding to relative abundance of each genus. Pathogenic genera (*Staphylococcus, Enterobacteriaceae, Pseudomonadaceae* and *Stenotrophomona*s) are displayed in variations of red, orange or purple color. In the case of *Pseudomonadaceae_unclassified*, *Enterobacteriaceae_unclassified* and *Pasterellaceae_unclassified*, classification to specific genera within these families was not accomplished, and thus we utilized family level descriptors for these genera.

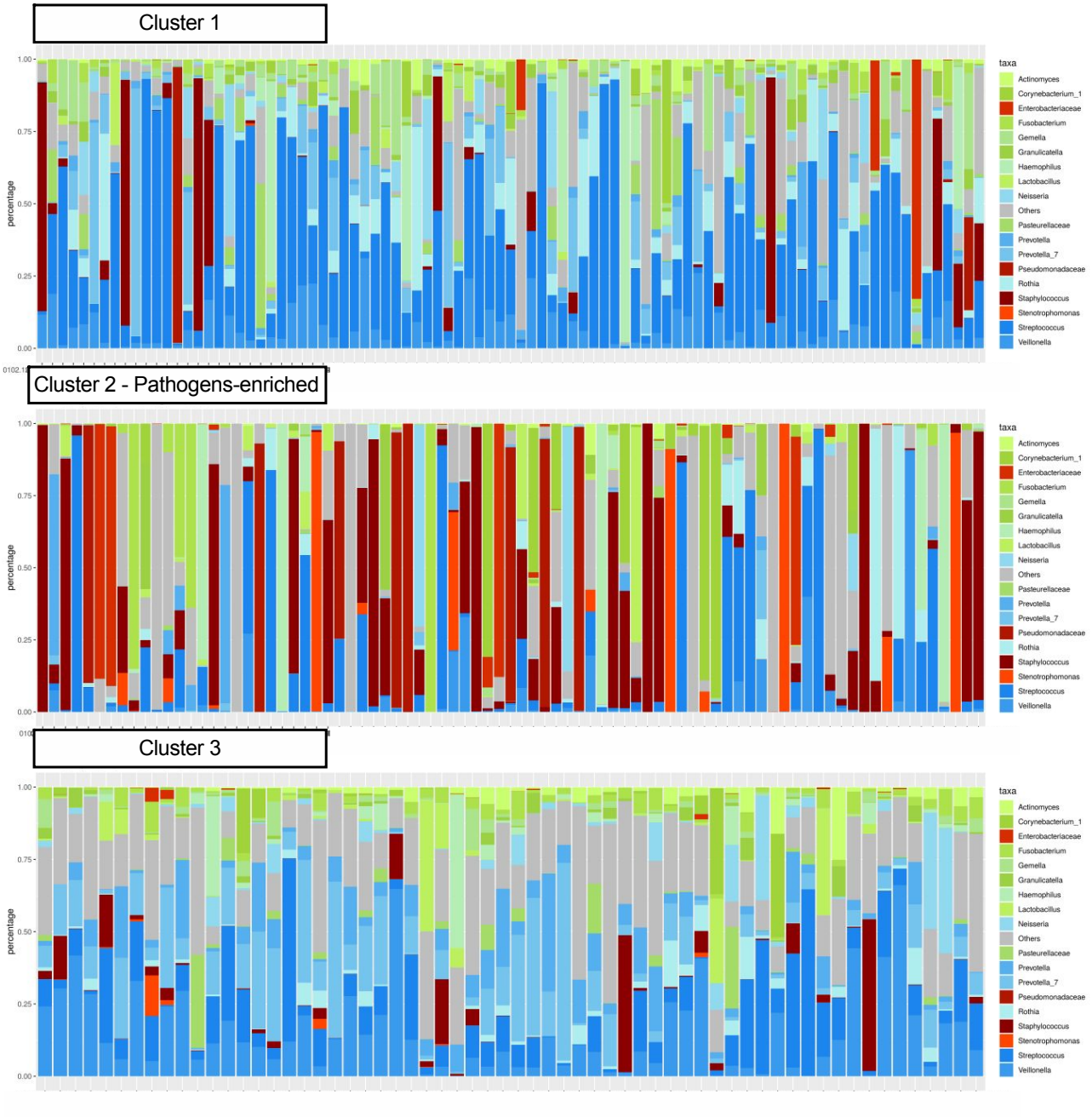
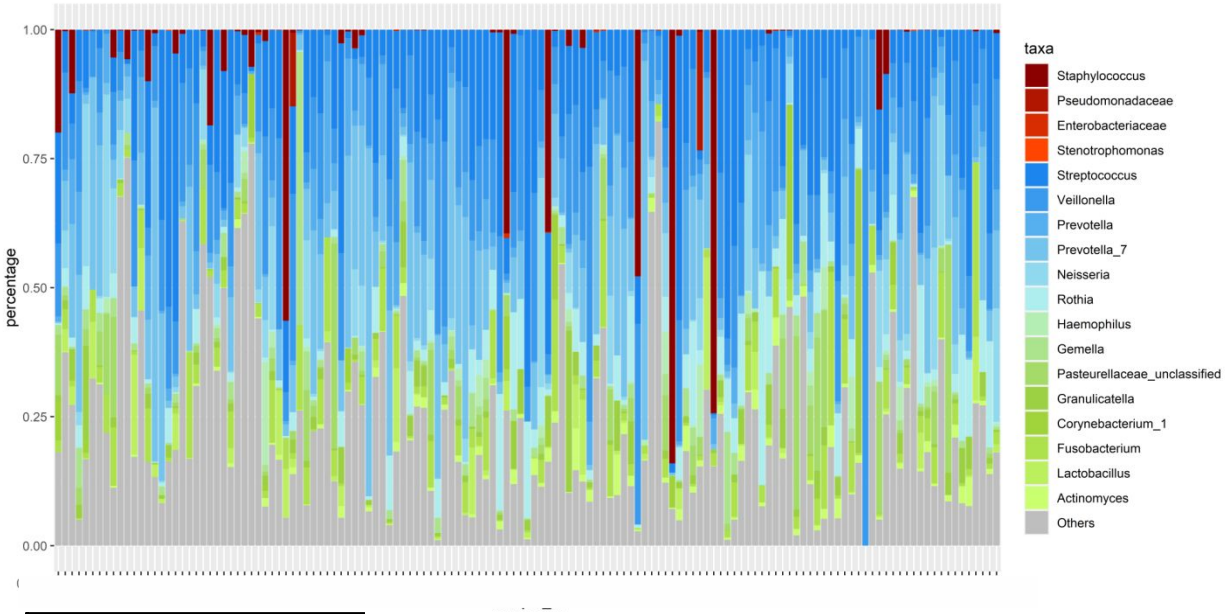

Cluster 1
Cluster 2 - Pathogens-enriched
Cluster 3
percentage
1.00
0.75
0.50
0.25
0.00
taxa
Actinomyces
Corynebacterium_1
Enterobacteriaceae
Fusobacterium
Gemella
Granulicatella
Haemophilus
Lactobacillus
Neisseria
Others
Pasteurellaceae
Prevotella
Prevotella_7
Pseudomonadaceae
Rothia
Staphylococcus
Stenotrophomonas
Streptococcus
Veillonella

Figure E7. Taxonomic bar plots for individual oral swab samples, stratified by Dirichlet Multinomial Models clusters. Taxonomic composition is shown as stacked bar-graphs, with each bar representing a patient's community, with taxa colored individually and heights of component bars corresponding to relative abundance of each genus. Pathogenic genera (*Staphylococcus, Enterobacteriacae, Pseudomonadacae* and *Stenotrophomona*s) are displayed in variations of red, orange or purple color. In the case of *Pseudomonadaceae_unclassified*, *Enterobacteriaceae_unclassified* and *Pasterellaceae_unclassified*, classification to specific genera within these families was not accomplished, and thus we utilized family level descriptors for these genera.

Cluster 1



Cluster 2 - Pathogens-enriched

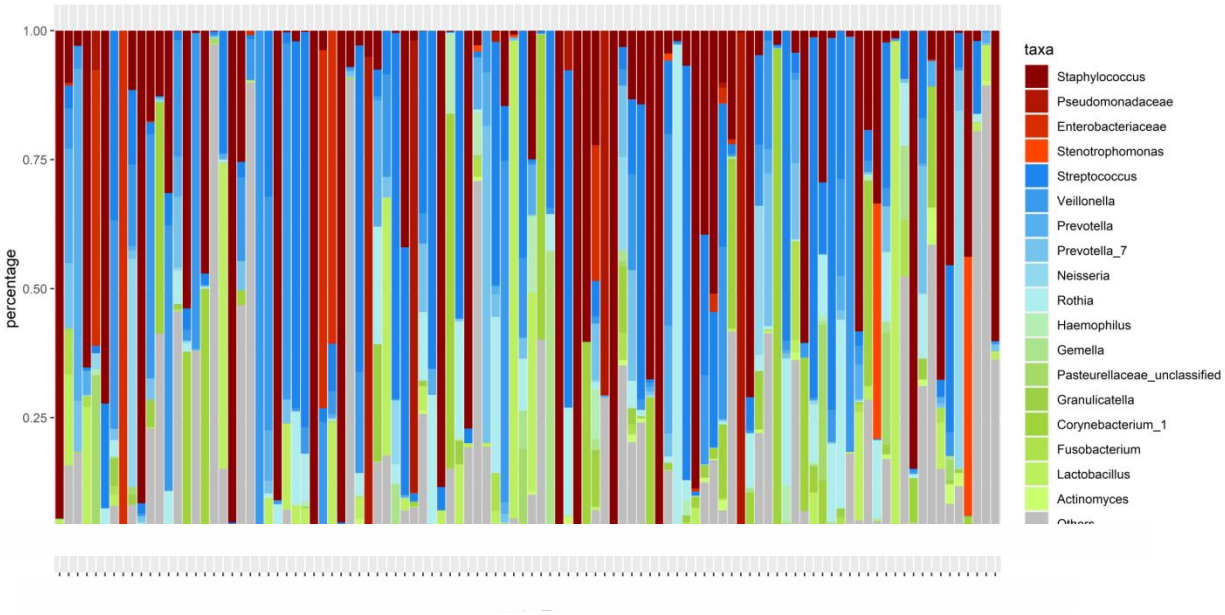


Figure E8: Relative abundance of individual genera is associated with clinical outcomes and host-response subphenotypes in patients without positive respiratory cultures. This analysis was restricted to patients for whom respiratory specimen cultures were reported as negative (no growth or only normal respiratory flora detected, n=152) as well as those for whom no cultures were obtained (n=76). This exploratory analysis serves to investigate individual genera associations in subjects for whom we had no clinically available information about the underlying microbial communities in lower respiratory tract samples, as was the case in subjects with positive respiratory specimen cultures. A. Endotracheal aspirate (ETA) genera. We examined for associations between additive log ratio transformed relative abundance for the top 10 genera in each cluster (total 18 unique genera) shown in y-axis with three outcome variables: Hyper-inflammatory subphenotype, 30-day mortality (logistic regression models) and Ventilator-Free Days (VFDs - linear regression model). Models were adjusted for age, COPD and antibiotic exposures. In each column, the direction of the effect size of the co-efficient and the statistical significance for each genus-outcome association are visually represented by color coding (protective effect shown in blue and adverse effect shown in red) and the size of each circle, respectively. *Enterobacteriaceae* relative abundance was associated with higher mortality, whereas typical members of the normal lung microbiome (e.g. *Prevotella_7* and *Streptococcus*) were associated with improved outcomes. A. Oral swab genera. Among the 14 unique genera examined in oral swabs, high relative abundance of *Staphylococcus* genera was associated with fewer VFDs, whereas high relative abundance of typical members of the normal lung microbiome (e.g. *Prevotella_7*, *Streptococcus, Veillonella, Rothia* etc.)

was associated with improved outcomes (mainly more VFDs). Associations that remained significant following adjustment for multiple testing with the Benjamini-Hochberg method are highlighted with asterisks (* - adjusted p<0.05). In the case of *Pseudomonadaceae_unclassified*, *Enterobacteriaceae_unclassified* and *Pasterellaceae_unclassified*, classification to specific genera within these families was not accomplished, and thus we utilized family level descriptors for these genera.

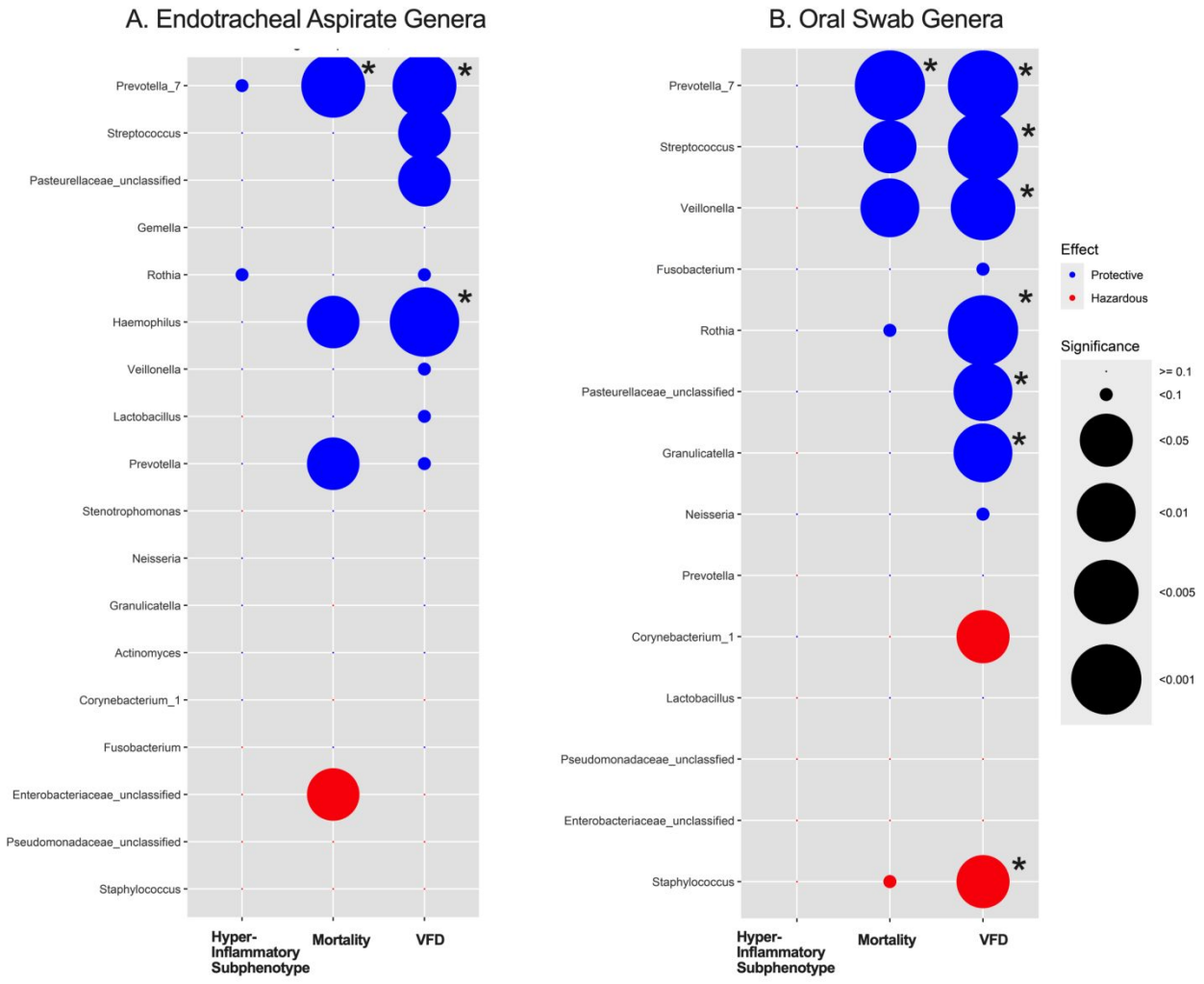
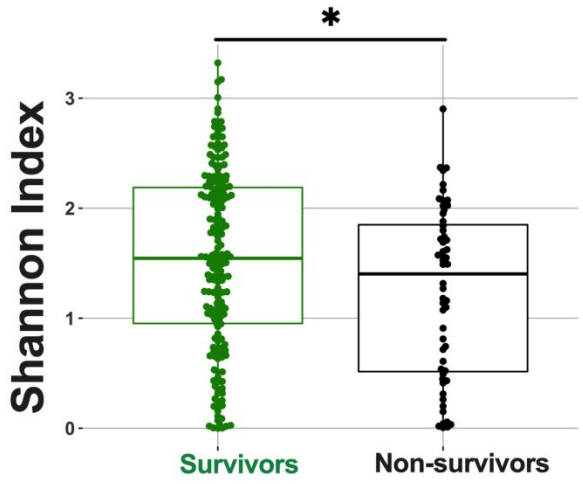
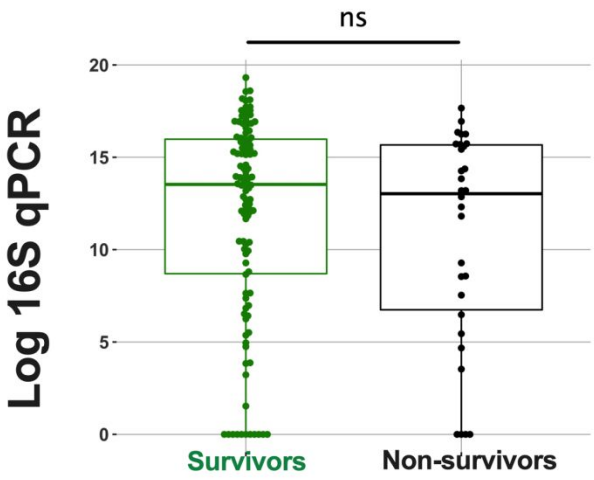


Figure E9: Supervised analysis for examination of associations between mortality and lung microbiota associations. Patients were stratified in survivors vs. non-survivors and then we conducted the following comparisons between groups: A. Alpha diversity, B. Bacterial load by quantitative PCR, C. Beta diversity by permutational analysis of variance of Manhattan distances and D. Differentially abundant genera (additive log ratio transformation of relative abundance) for the 18 most abundant genera in endotracheal aspirate samples. Raw and adjusted for multiple testing (Benjamini-Hochberg method) p-values are shown. Genera enriched in survivors are shown in blue (“protective”) whereas genera enriched in non-survivors are shown in red (“hazardous” genera).

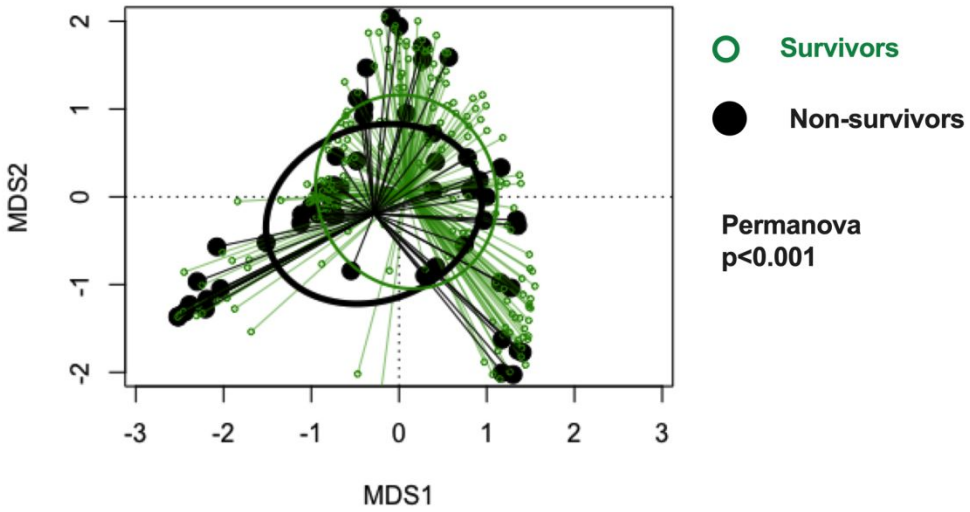
A. Alpha Diversity



B. Bacterial Load



C. Beta Diversity

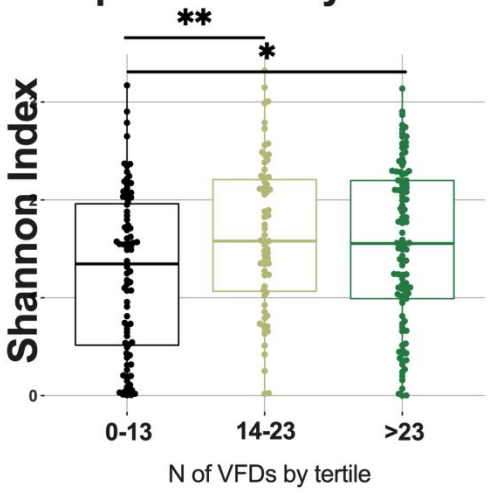


D. Differentially abundant genera

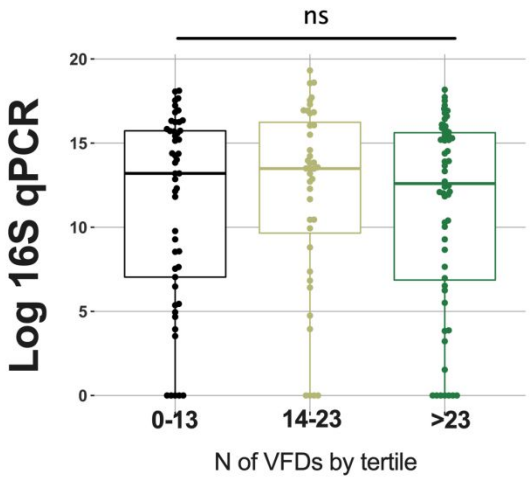
Survivors			Non-survivors		
Enriched Genera	Raw p-value	Adjusted p-value	Enriched Genera	Raw p-value	Adjusted p-value
Prevotella_7	**0.01**	0.07	**Staphylococcus**	**0.03**	0.11
			Pseudomonodacae	**0.03**	0.11
			Escherichia	**0.01**	0.07
			Enterobacteriaceae	**0.007**	0.07

Figure E10: Supervised analysis for examination of associations between ventilator-free days (VFDs) and lung microbiota associations. Patients were stratified in three tertiles of VFDs (0-13, 14-23, >23) and then we conducted the following comparisons between groups: A. Alpha diversity, B. Bacterial load by quantitative PCR, C. Beta diversity by permutational analysis of variance of Manhattan distances and D. Differentially abundant genera (additive log ratio transformation of relative abundance) for the 18 most abundant genera in endotracheal aspirate samples. Raw and adjusted for multiple testing (Benjamini-Hochberg method) p-values are shown. Genera enriched in patients with more VFDs are shown in blue (“protective”) whereas genera enriched in patients with fewer VFDs are shown in red (“hazardous” genera).

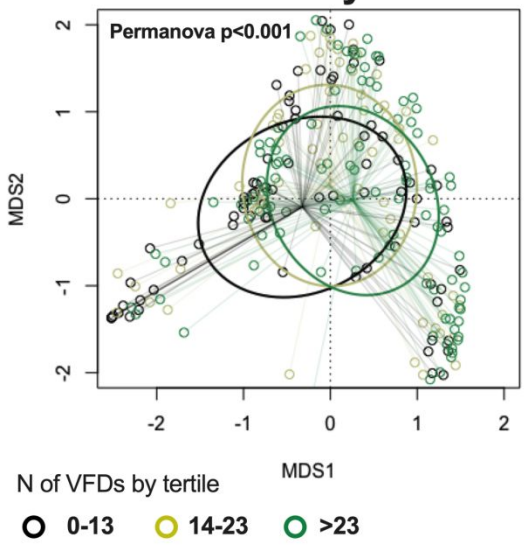
A. Alpha Diversity



B. Bacterial Load



C. Beta Diversity

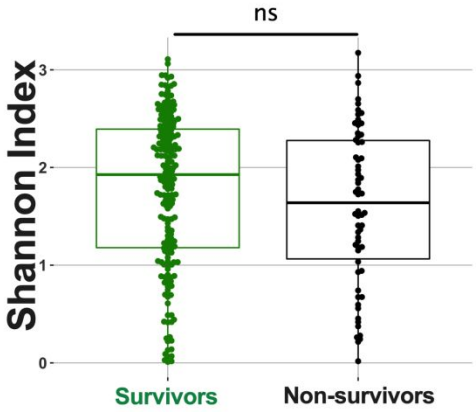


D. Differentially abundant genera by VFD tertiles

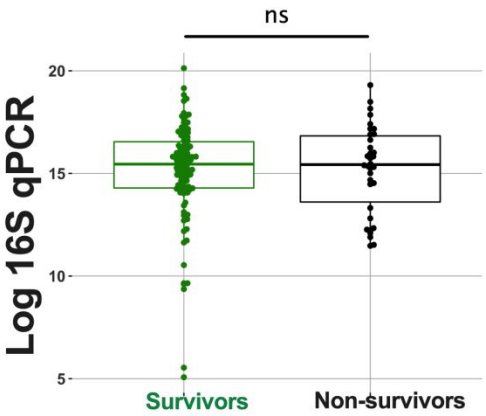
Kruskal Wallis Test between 3 groups		
Enriched Genera	Raw p-value	Adjusted p-value
Staphylococcus	**0.03**	0.17
Streptococcus	**0.002**	**0.03**
Escherichia	**0.04**	0.17

Figure E11: Supervised analysis for examination of associations between mortality and oral swab microbiota associations. Patients were stratified in survivors vs. non-survivors and then we conducted the following comparisons between groups: A. Alpha diversity, B. Bacterial load by quantitative PCR, C. Beta diversity by permutational analysis of variance of Manhattan distances and D. Differentially abundant genera (additive log ratio transformation of relative abundance) for the 14 most abundant genera in oral swab samples. Raw and adjusted for multiple testing (Benjamini-Hochberg method) p-values are shown. Genera enriched in survivors are shown in blue ("protective).

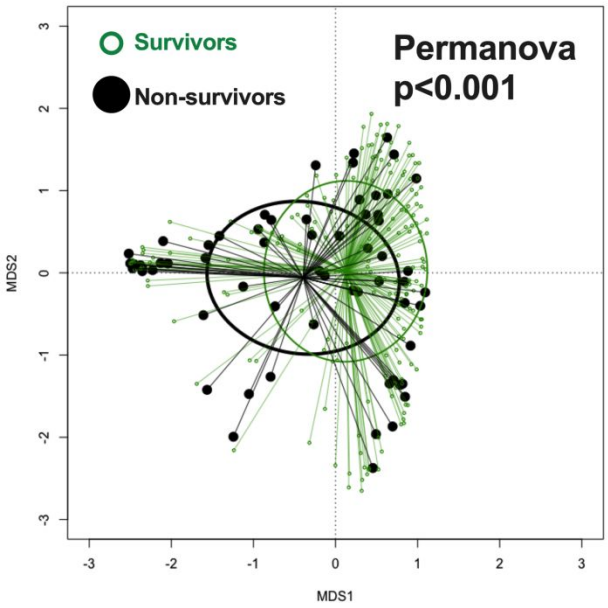
A. Alpha Diversity



B. Bacterial Load



C. Beta Diversity

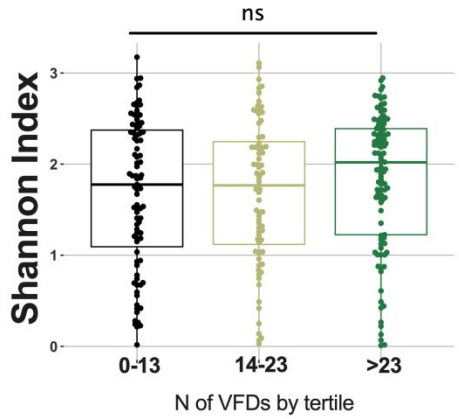


D. Differentially abundant genera

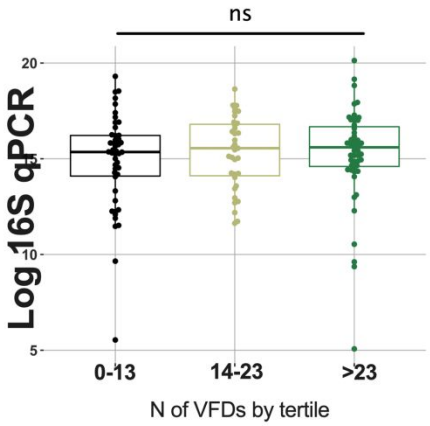
Survivors		
Enriched Genera	Raw p-value	Adjusted p-value
Prevotella_7	**0.0001**	**0.002**

Figure E12: Supervised analysis for examination of associations between ventilator-free days (VFDs) and oral swab microbiota associations. Patients were stratified in three tertiles of VFDs (0-13, 14-23, >23) and then we conducted the following comparisons between groups: A. Alpha diversity, B. Bacterial load by quantitative PCR, C. Beta diversity by permutational analysis of variance of Manhattan distances and D. Differentially abundant genera (additive log ratio transformation of relative abundance) for the 18 most abundant genera in oral swab samples. Raw and adjusted for multiple testing (Benjamini-Hochberg method) p-values are shown. Genera enriched in patients with more VFDs are shown in blue (“protective”) whereas genera enriched in patients with fewer VFDs are shown in red (“hazardous” genera).

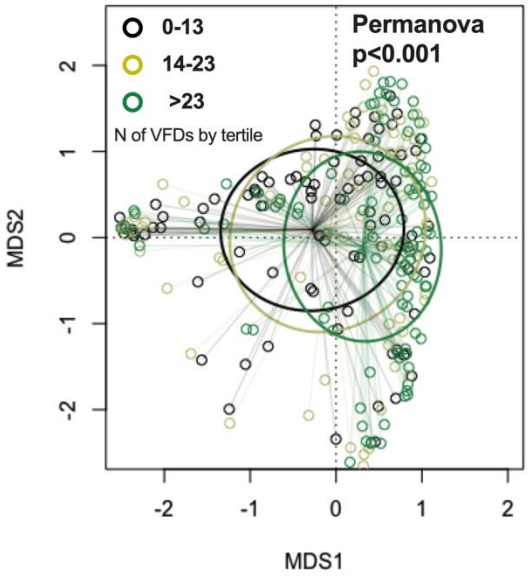
A. Alpha Diversity



B. Bacterial Load



C. Beta Diversity



D. Differentially abundant genera

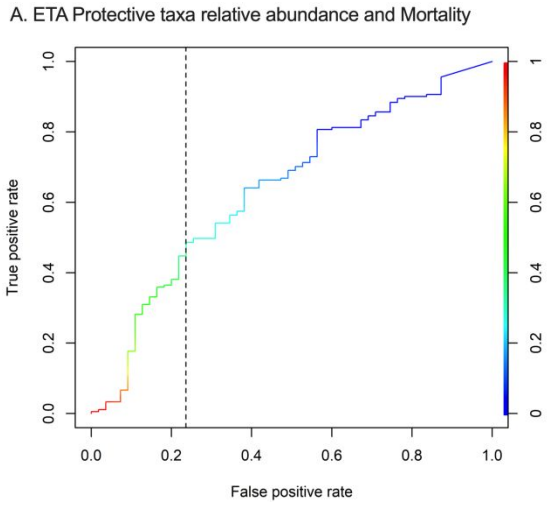
Kruskal Wallis Test between 3 groups		
Enriched Genera	Raw p-value	Adjusted p-value
Streptococcus	**0.0007**	**0.003**
Prevotella_7	**0.0002**	**0.0003**
Rothia	**0.001**	**0.004**
Neisseria	**0.004**	**0.009**
Veillonella	**0.02**	**0.03**
Granulicatella	**0.002**	**0.006**
Pasteurellaceae	**0.0001**	**0.0007**
Enterobacteriaceae	**0.03**	0.06
Pseudomonadaceae	**0.05**	0.07

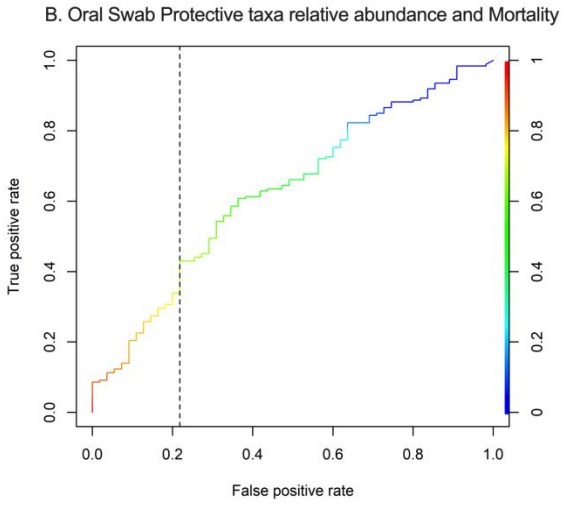
Figure E13: Receiver operator characteristic (ROC) curve analyses for derivation of thresholds of protective genera relative abundance and the creation of a dysbiosis index. In all these analyses, we considered specificity of thresholds as more clinically relevant than maximum sensitivity, and therefore we did not utilize mathematically optimal solutions for threshold determination but examined for thresholds of protective genera relative abundance in the range of a false positive rate of 20% (to ensure an acceptable specificity of ~80%). As protective genera in ETA samples, we considered all genera with at least a nominal association ($p<0.1$) with improved outcomes on Figure 4. This resulted in a set of 5 genera: *Prevotella_7, Streptococcus, Gemella, Rothia* and *Haemophilus*. We considered the relative abundance for each of these genera as a continuous variable and examined for association with mortality in the ROC curve of panel A. A threshold of <30% offered an AUC of 0.63 for mortality prediction with a false positive rate of 24%. Therefore, we considered a threshold of ≥30% for ETA protective genera relative abundance to be indicative of communities without dysbiosis. For oral swabs, we considered all genera with at least a nominal association ($p<0.1$) with improved outcomes on Figure 4. This resulted in a set of 7 genera: *Prevotella_7, Neisseria, Streptococcus, Granulicatella, Rothia, Veillonella* and *Pasteurellaceae*. A threshold of <70% offered an AUC of 0.63 for mortality prediction with a false positive rate of 22% in panel B. Therefore, we considered a threshold of ≥70% for ETA protective genera relative abundance to be indicative of communities without dysbiosis. For alpha diversity, we defined that a Shannon index ≥1.98 would be indicative of dysbiosis, based on the distribution of the Shannon index in the Dirichlet Multinomial Models (DMM) clusters, i.e. a Shannon index

of ≥1.98 discriminated the good prognosis cluster 3 for ETA and cluster 1 for oral swabs from the rest of the cohort.

Lastly, we combined the thresholds of Shannon index and protective genera relative abundance into a Dysbiosis Index for ETA and oral swabs.

- ETA Dysbiosis Index:
    - Shannon index ≥1.98 and ETA protective genera relative abundance ≥30%: No Dysbiosis
    - Shannon index <1.98 and/or ETA protective genera relative abundance <30%: Dysbiosis
- Oral Swab Dysbiosis Index:
    - Shannon index ≥1.98 and oral swab protective genera relative abundance ≥70%: No Dysbiosis
    - Shannon index <1.98 and/or oral swab protective genera relative abundance <70%: Dysbiosis





Bibliography

1. Kitsios GD, Fitch A, Manatakis DV, Rapport SF, Li K, Qin S, Huwe J, Zhang Y, Doi Y, Evankovich J, Bain W, Lee JS, Methé B, Benos PV, Morris A, McVerry BJ. Respiratory microbiome profiling for etiologic diagnosis of pneumonia in mechanically ventilated patients. *Front Microbiol* 2018;9:1413.
2. Raith EP, Udy AA, Bailey M, McGloughlin S, MacIsaac C, Bellomo R, Pilcher DV, Australian and New Zealand Intensive Care Society (ANZICS) Centre for Outcomes and Resource Evaluation (CORE). Prognostic Accuracy of the SOFA Score, SIRS Criteria, and qSOFA Score for In-Hospital Mortality Among Adults With Suspected Infection Admitted to the Intensive Care Unit. *JAMA* 2017;317:290–300.
3. Gajic O, Dabbagh O, Park PK, Adesanya A, Chang SY, Hou P, Anderson H, Hoth JJ, Mikkelsen ME, Gentile NT, Gong MN, Talmor D, Bajwa E, Watkins TR, Festic E, Yilmaz M, Iscimen R, Kaufman DA, Esper AM, Sadikot R, Douglas I, Sevransky J, Malinchoc M, U.S. Critical Illness and Injury Trials Group: Lung Injury Prevention Study Investigators (USCIITG-LIPS). Early identification of patients at risk of acute lung injury: evaluation of lung injury prediction score in a multicenter cohort study. *Am J Respir Crit Care Med* 2011;183:462–470.
4. Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L, Jablonski K, Kleerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman A, Schmidt TM, Weinstock GM, Lung HIV Microbiome Project. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. *Am J Respir Crit Care Med* 2013;187:1067–1075.
5. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 2012;6:1621–1624.
6. Liu CM, Aziz M, Kachur S, Hsueh P-R, Huang Y-T, Keim P, Price LB. BactQuant: an enhanced broad-coverage bacterial quantitative real-time PCR assay. *BMC Microbiol* 2012;12:56.
7. Morgulis A, Gertz EM, Schäffer AA, Agarwala R. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. *J Comput Biol* 2006;13:1028–1040.
8. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E, Methé B, DeSantis TZ, Human Microbiome Consortium, Petrosino JF, Knight R, Birren BW. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 2011;21:494–504.
9. Kitsios GD, Yang L, Manatakis DV, Nouraie M, Evankovich J, Bain W, Dunlap DG, Shah F, Barbash IJ, Rapport SF, Zhang Y, DeSensi RS, Weathington NM, Chen BB, Ray P, Mallampalli RK, Benos PV, Lee JS, Morris A, McVerry BJ. Host-Response Subphenotypes Offer Prognostic Enrichment in Patients With or at Risk for Acute Respiratory Distress Syndrome. *Crit Care Med* 2019;47:1724–1734.
10. Holmes I, Harris K, Quince C. Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS One* 2012;7:e30126.